

بسم الله الرحمن الرحيم

دانشگاه علم و صنعت ایران پروژه درس مبانی بینایی کامپیوتر

یکی از پر کاربرد ترین مسائل حوزه بینایی کامپیوتر، مسائل مربوط به OCR است. هدف این پروژه، پیاده سازی یکی از مسائل مربوط به این حوزه است.

برای انجام این پروژه، ابتدا شما باید تعدادی تصویر از کارت بانکی و کارت ملی های مختلف (حداقل تعداد 10 کارت) در زوایای مختلف تهیه کنید و مانند شکل زیر با تکنیک های پردازش تصویر محدوده مربوط به کارت را برش بزنید. سپس، محتوای مورد نظر (شماره کارت بانکی یا شماره ملی) از کارت را بخوانید.



این پروژه می تواند به فازهای زیر شکسته شود که برای انجام آن می توانید از تمام ابزارهایی که در درس خواندید استفاده کنید (در قسمت هایی که از یادگیری عمیق استفاده می کنید، تهیه دیتاست بزرگ بسیار دشوار است اما می توان دیتاست ساختگی با فونت ها و شرایط گوناگون به تعداد بسیار زیاد و به سادگی تهیه کرد):

1. برش بخش مربوط به کارت در تصویر.
 2. تشخیص نوع کارت (کارت بانکی یا کارت ملی).
 3. تعیین محل مربوط به اعداد روی کارت.
 4. تعیین فارسی یا انگلیسی بودن اعداد.
 5. پیش بینی اعداد. دو روش مختلف برای این بخش قابل انجام است. یکی را به دلخواه انتخاب کنید:
- یک فایل شامل ده تصویر ground truth از اعداد 0 تا 9 ایجاد کنید. حال برای 10 مورد با استفاده از template matching و ابزارهای دیگری که در درس خواندید (ابزارهای رفع نویز و ...) عمل پیش بینی را انجام دهید. (دقت شود که الگوریتم شما باید دقت مناسبی داشته باشد)
 - یک مدل یادگیری عمیق ساخته و آن را یک بار با اعداد فارسی و بار دیگر با اعداد انگلیسی آموزش داده و عمل پیش بینی را با استفاده از تابعی که نوشته شده روی تصاویر خود انجام دهید.

بخش امتیازی: تشخیص تاریخ انقضاء و CVV2 کارت های بانکی نیز عمل پیش بینی را انجام دهید (امتیازی).

در پایان نتایج به دست آمده را مقایسه و تحلیل کنید. لازم به ذکر است میزان دقت در نتایج به دست آمده بسیار مهم است.

توجه شود که از هیچ کتابخانه‌ای غیر از کتابخانه‌های **import** شده نمی‌توانید استفاده کنید.

در انجام این پروژه مجاز به استفاده از تمام ابزارهای خوانده شده در درس هستید. همچنین، در صورت استفاده از ابزارهایی که در درس مطرح نشده است لازم است به جزئیات آن مسلط باشید. اجرای پروژه در قالب گروه‌های دو نفره خواهد بود. لطفاً حداکثر تا تاریخ 15 اردیبهشت اسامی اعضای تیم خود را به آقای چینی‌فروشان اطلاع دهید.

خروجی کار شما عبارت است از:

- یک گزارش کامل از تمام کارهایی که در این پروژه انجام داده‌اید و نتایجی که بدست آورده‌اید. در نوشتن گزارش توجه داشته باشید که به هر مرجعی که استفاده می‌کنید (چه از لحاظ تئوری، چه از لحاظ کدنویسی) به دقت ارجاع بدهید.
- یک کد کامل که بتواند بر روی سیستم اجرا شود و نتایج را بر روی یک مجموعه داده جدید ذخیره کند. این کد باید شامل یک تابع با نام `predict` باشد که تصویر و طرح مربوطه را دریافت کند و مختصات ترک‌های داخل آن را بازگرداند. این تابع در یک کد دیگر `Import` خواهد شد و بر روی یک مجموعه از تصاویر جدید (با طرح‌های متفاوت) ارزیابی خواهد شد. نتایج کار شما از لحاظ کمی (دقت و سرعت) و کیفی با نتایج دیگران مقایسه خواهد شد.

موفق باشید.